ORIGINAL ARTICLE

Comparison of STD prevalences in the Mwanza, Rakai, and Masaka trial populations: the role of selection bias and diagnostic errors

K K Orroth, E L Korenromp, R G White, J Changalucha, S J de Vlas, R H Gray,
P Hughes, A Kamali, A Ojwiya, D Serwadda, M J Wawer, R J Hayes, H Grosskurth

....

Objectives: To assess bias in estimates of STD prevalence in population based surveys resulting from diagnostic error and selection bias. To evaluate the effects of such biases on STD prevalence estimates from three community randomised trials of STD treatment for HIV prevention in Masaka and Rakai, Uganda and Mwanza, Tanzania.

Methods: Age and sex stratified prevalences of gonorrhoea, chlamydia, syphilis, HSV-2 infection, and trichomoniasis observed at baseline in the three trials were adjusted for sensitivity and specificity of diagnostic tests and for sample selection criteria.

Results: STD prevalences were underestimated in all three populations because of diagnostic errors and selection bias. After adjustment, gonorrhoea prevalence was higher in men and women in Mwanza (2.8% and 2.3%) compared to Rakai (1.1% and 1.9%) and Masaka (0.9% and 1.8%). Chlamydia prevalence was higher in women in Mwanza (13.0%) compared to Rakai (3.2%) and Masaka (1.6%) but similar in men (2.3% in Mwanza, 2.7% in Rakai, and 2.2% in Masaka). Prevalence of trichomoniasis was higher in women in Mwanza compared to women in Rakai (41.9% versus 30.8%). Herpes simplex virus type 2 (HSV-2) seroprevalence and prevalence of serological syphilis (TPHA+/RPR+) were similar in the three populations but the prevalence of high titre syphilis (TPHA+/RPR ≥1:8) in men and women was higher in Mwanza (5.6% and 6.3%) than in Rakai (2.3% and 1.4%) and Masaka (1.2% and 0.7%).

Conclusions: Limited sensitivity of diagnostic and screening tests led to underestimation of STD prevalence in all three trials but especially in Mwanza. Adjusted prevalences of curable STD were higher in Mwanza than in Rakai and Masaka.

See end of article for
authors' affiliations
. . . .

Correspondence to:
Kate K Orroth, MPH,
Infectious Disease
Epidemiology Unit, London
School of Hygiene and
Tropical Medicine, Keppel
Street, London WC1E 7HT,
UK;
kate.orroth@lshtm.ac.uk

Accepted for publication
2 September 2002
. . . .

Prevalences of sexually transmitted diseases (STD) are often compared between populations in order to gain a better understanding of STD and HIV epidemiology or to determine which control strategy may be most effective in a given epidemiological situation. Three community randomised trials of STD treatment as HIV prevention strategies have been conducted in east Africa in Mwanza, Tanzania, and Masaka and Rakai, Uganda.[1–3] In the Mwanza trial, improved STD case management was associated with a 38% reduction in HIV incidence.[4] In contrast, in Uganda a trial of information, education, and communication coupled with improved STD case management had no impact on HIV incidence in Masaka district[5] and STD mass treatment had no impact on HIV incidence in Rakai district.[3] Baseline prevalence of gonorrhoea and chlamydia was about 5% or less in the populations while that for serological syphilis was 10–15%. Various hypotheses have been suggested to explain the apparently contrasting results, including that the Mwanza and Rakai trials differed with respect to the stage of the HIV epidemic, that the proportion of ulcers due to herpes simplex virus type 2 (HSV-2) was higher in Rakai than Mwanza, and that differences in the interventions—that is, continuous treatment of symptomatic STD (as in Mwanza) may have a larger effect than treating all STD periodically (as in Rakai).[6] Another potential explanation is that the prevalence of curable STD differed between the three sites. Based on the reduction in HIV incidence in Mwanza and the association between STD and HIV transmission we would expect higher prevalence of STD in Mwanza than in Masaka and Rakai. However, the STD prevalences reported in the three trials cannot be directly compared, as different diagnostic tests and sampling strategies were used.

STD prevalences observed in surveys, such as in the east African trials, depend upon selection of the sample for measurement, extrapolation of the prevalence measured in this sample to the general population and the diagnostic technique used. Selection of the population in which the STD is measured may bias observed STD prevalence either upward or downward. The net effect of diagnostic performance depends on the true prevalence level: if prevalence is low, specificity is most important but if prevalence is high sensitivity is critical. Diagnostic errors can lead to either overestimation or underestimation of prevalence.

The measurement errors in STD prevalences due to screening and test diagnostics were assessed for the Mwanza, Masaka, and Rakai trials. Sampling bias was also considered when comparing measured prevalences in the three populations. We adjusted the observed baseline prevalences of *Neisseria gonorrhoeae* (NG), *Chlamydia trachomatis* (CT), *Trichomonas vaginalis* (TV), syphilis, and HSV-2 infection for these biases, and then compared prevalences between the three populations. We also considered the implications of these comparisons for the interpretation of the discrepant trial outcomes.

METHODS

Prevalences by 5 year age groups for men and women were adjusted for screening and diagnostic techniques used and then for selection bias, as explained below. After adjustment, the overall prevalence of each infection in adults aged 15–54 years was obtained after standardising for age using the average population structure of the three populations based on

census data.[7][8] When data were not available for the entire age range 15–54, restricted age ranges were used.

Diagnostic adjustments

Observed prevalences were adjusted for both the sensitivity and specificity of the diagnostic tests according to equation 1[9]:

$$p_{true} = \frac{(p_{observed} + Sp - 1)}{(Se + Sp - 1)}$$

where, p_{true} = true prevalence, $p_{observed}$ = observed prevalence, Se = sensitivity, and Sp = specificity.

For outcomes where diagnostics were performed on only a subset of the population based on results of a screening test, prevalences were also adjusted for the screening algorithm. The formula for this combined adjustment is:

$$P_{true} = \frac{p_{observed} - p_{screen}(1 - Sp)}{Se'(Se + Sp - 1)}$$

where p_{true} = true prevalence, $p_{observed}$ = observed prevalence assuming those negative on the screening test to be negative, p_{screen} = prevalence of a positive screening test, Se = sensitivity of the diagnostic test, Sp = specificity of the diagnostic test, and Se′ = sensitivity of the screening test. The derivation of equation 2 is given in the appendix.

Estimates of sensitivity and specificity for diagnostic tests

The sensitivity and specificity of the tests used were estimated based on values documented in the literature for populations from all geographical locations. Initially eligible studies were limited to those conducted in asymptomatic populations. However, for some of the tests estimates were only available for symptomatic populations so this criterion was dropped. When multiple studies were available we averaged test sensitivities and specificities across published studies, and the mean values were used for adjustment of the trial observations. To consider uncertainty in diagnostic test performance we used the highest documented value for sensitivity from any single study combined with the lowest value for specificity from any single study, to estimate a lower limit for the adjusted prevalence. Similarly, the upper limit was determined by using the lowest sensitivity and highest specificity.

To obtain estimates of sensitivity and specificity, we included only published studies in which the techniques used in the trials were directly compared in the same population using the same gold standard. These studies typically used an "expanded" gold standard in which discrepant results were resolved using another confirmatory test. Studies using culture alone as the gold standard were not used because the imperfect sensitivity of culture would incorrectly suggest a low specificity for the tests under study which have higher sensitivity than culture, such as LCR.[10]

Gonorrhoea and chlamydia

In Rakai and Masaka, ligase chain reaction (LCR, Abbott Laboratories, Abbot Park, IL, USA) on urine was used to diagnose NG and CT in men and women. In Mwanza, Gram stain on urethral smears was used to diagnose NG in men. For women, culture from endocervical swabs were used to diagnose NG. In Mwanza, enzyme immunoassay (EIA, IDEIA Chlamydia; Novo Nordisk Diagnostika, Cambridge, UK) from urethral swabs was used to diagnose CT in men, and EIA from endocervical swabs was used for CT in women.

In Mwanza, a screening test was used to select the sample of men to be tested for NG and CT with testing confined to

men who either tested positive on urine leucocyte esterase dipstick (LED; Nephur-Test + Leuco, Boehringer-Mannheim, Lewes, UK) tests, who complained of discharge during the interview, or who had discharge upon clinical examination. To allow adjustment for the performance of the screening test (according to equation 2), we estimated its sensitivity based on findings in a separate study conducted in 1996 in a rural community in Mwanza Region[11] (see below).

Trichomoniasis

The prevalence of *Trichomonas vaginalis* in women was assessed via InPouch TV culture (BioMed Diagnostics, San Jose, CA, USA) on self collected vaginal swabs in Rakai and by wet mount microscopy of vaginal smears in Mwanza. TV prevalence was not available for men in the three trials or from women in Masaka. In the reviewed literature, studies were included which directly compared the performance of InPouch TV culture to wet mount microscopy with either another culture method or polymerase chain reaction (PCR) to resolve discrepant results.

Syphilis

A non-treponemal test, the toluidine red unheated serum test (TRUST; New Horizons, Columbia, MD, USA), was used to screen the Rakai population for syphilis. Those testing positive on the TRUST test were further tested using the *Treponema pallidum* haemoagglutination assay (TPHA; TPHA Sera-Tek, Fujirebio, Tokyo, Japan) test.[12] In Masaka and Mwanza, the TPHA test was conducted on all study participants. A non-treponemal rapid plasma reagin (RPR; VD-25 Murex, Dartford, UK) test was conducted for those who were TPHA positive.[13] We compared the prevalence of serological syphilis (TPHA positive with any RPR/TRUST titre) and the prevalence of active, high titre syphilis (TPHA positive with RPR titre ≥1:8) in all three populations. The diagnostic performance of the RPR and TRUST tests are similar with 98% sensitivity and 99% specificity[14][15] and TPHA tests were used in both populations so no adjustment was needed to validly compare serological or active high titre syphilis prevalence as measured in the three sites.

HSV-2 seroprevalence

HSV-2 serology for the Rakai trial was conducted at the Centers for Disease Control (CDC) using an immunoblot assay which discriminates between antibodies for HSV-1 and HSV-2.[3] The diagnostic adjustments proposed for the Rakai data are based on comparison with another western blot assay as the gold standard.[16] Seroprevalence in Masaka and Mwanza was measured using a monoclonal antibody blocking immunoassay test and the diagnostic adjustment is based on the comparison of this test with the aforementioned western blot in a rural African population.[17]

Estimation of selection bias

If random samples of the trial population were selected or the entire trial population was evaluated, we adjusted STD prevalences for diagnostic test performance only. This was the case for NG, CT, TV, and syphilis in Rakai; for NG, CT, HSV-2 and syphilis in Masaka; and for HSV-2 and syphilis in Mwanza.[2][12][13][18] We restricted comparison of STD prevalences to the sex and age ranges for which data were available from all sites. Thus, comparisons for NG and CT were restricted to women and men aged 15–39 years. Comparisons for TV included women aged 15–49 years. We compared HSV-2 seroprevalence in 15–29 year olds and syphilis in 15–54 year olds.

Where only non-random samples of the trial populations had been measured, additional adjustments were considered. In women in Mwanza, prevalences of NG, CT, and TV were measured only among antenatal clinic (ANC) attendees. ANC data from Mwanza town have been shown to underestimate

HIV prevalence and syphilis prevalence in women in the general population.[19][20] Similarly, unpublished analysis of Rakai trial data found age standardised NG prevalence among 15–39 year olds to be 0.7% among self reported pregnant women and 1.6% among women in the general population. For CT in Rakai, the corresponding prevalences were 2.9% and 2.7% and for TV, prevalences were 24% in both groups (Eline Korenromp, personal communication). These data suggest levels of curable STD are similar or slightly lower in pregnant women compared to the general population in some African populations. However, it is unclear if this can be assumed to apply to women attending antenatal clinics in rural Mwanza. It may be that selection biases associated with antenatal populations differ between urban and rural areas, or that HIV associated effects on fertility may lead to different biases in rural Mwanza (4% HIV prevalence) and Rakai (16% HIV prevalence). Therefore, the Mwanza data were not adjusted for selection of the ANC sample.

HSV-2 seroprevalence in Rakai was measured only among individuals reporting they were sexually active. Hence the population adjustment must take into account the proportion sexually active in the different age strata in the sampled group (15–29 year olds). Diagnostic adjustments were applied to the observed prevalence in each age group. Then this adjusted prevalence was multiplied by the proportion sexually active for each age group and finally the prevalence was age standardised.

Sensitivity analysis
A sensitivity analysis was conducted to ascertain the relative importance of the different adjustments performed to validly compare STD prevalence across the sites. We used male NG prevalence and female HSV-2 prevalence as examples for the sensitivity analysis. To determine the relative importance of the sensitivity and specificity of diagnostic tests we varied each alone and then the combination. The sensitivity and prevalence of a positive screening test were varied one at a time, together and in combination with diagnostic test performance. For screening test performance, we used the values observed for NG in men during the Misungwi study and lower and upper limits. The lower bound for the sensitivity of the screening test was 44% (half that of the default value) and the upper bound was 95%. To assess the importance of selection bias we varied the proportion sexually active in the Rakai HSV-2 data. The proportion reporting sexual activity was used as the base case adjustment. To estimate an upper limit for HSV-2 prevalence, we assumed underreporting of sexual activity was 50% (of those who denied sexual activity, half were truly sexually active) among 15–19 year olds, 5% among 20–24 year olds, and 0% in 25–29 year olds.

RESULTS
Table 1 shows published estimates of the sensitivity and specificity of the diagnostic tests. Where possible, the average sensitivity and specificity for each test are also given. For NG in men, only one study directly comparing the diagnostic properties of LCR on urine to Gram stain of urethral smears was identified for an STD clinic population and so we used data from this study.[21] Similarly, only one study was available to determine the diagnostic performance of each of the HSV-2 serological tests. For the tests used in Mwanza, sensitivity was generally low, mean estimates ranging between 49% and 91%. For the Masaka tests, mean sensitivities ranged between 84% and 91%. For the Rakai tests, mean sensitivities ranged between 79% and 92%. Mean specificities were between 93% and 100% for Mwanza and Masaka and were 100% for all tests used in Rakai.

Table 2 lists all adjustments required for the STD prevalence comparison. For diagnostic tests, these were based on the means and ranges given in table 1. For the screening test used for NG and CT testing in men in Mwanza, we estimated sensitivity at 53% for CT and 88% for NG; prevalence of a positive screening test, needed for the adjustment according to equation 2 was about 25% for all age groups. For selection bias adjustment of the Rakai HSV-2 data, measured proportions sexually active for each age group are listed.

Table 3 shows observed prevalences with 95% confidence intervals and adjusted prevalences of CT, NG, TV, and HSV-2 serology. For serological (TPHA+/RPR+) and active high titre (TPHA+/RPR⩾1:8) syphilis only the directly observed prevalences are given as adjustment was not necessary. In general, similar sample sizes were used in the three studies so sampling errors of the observed prevalences were similar across the three trials.

In women, observed prevalences of CT and NG were higher in Mwanza than in Rakai and Masaka. After adjustment for diagnostics these differences were accentuated. Observed and adjusted prevalence of CT was higher in Rakai than in Masaka but prevalence of NG was similar in the two Ugandan sites. Observed TV prevalence was similar in Mwanza and Rakai, but after adjustment, prevalence was higher in Mwanza compared to Rakai. Observed prevalence of serological syphilis was higher in Rakai and Masaka than in Mwanza but prevalence of active high titre syphilis was much higher in Mwanza than in Rakai and Masaka. Also for HSV-2, seroprevalence was slightly lower in the Ugandan sites compared to Mwanza even after allowing for possible under-reporting of sexual activity in Rakai.

In men, observed CT prevalence was higher in Rakai and Masaka than in Mwanza but after adjustment there was little difference between sites. Observed NG prevalence was higher in Mwanza than the Ugandan sites and the difference was accentuated after adjustment. Similar to results for women, the observed prevalence of serological syphilis was higher in Masaka and Rakai than in Mwanza but prevalence of active high titre syphilis was much higher in Mwanza than in Masaka and Rakai. Levels of HSV-2 were higher in Rakai and Masaka compared to Mwanza.

The reported ranges indicate that the outcomes were fairly robust against uncertainty in our estimates of sensitivity and specificity of the diagnostic tests. Exceptions were CT and TV in females in Mwanza, for which estimated prevalences changed considerably when replacing the mean estimates for sensitivity by values documented from single studies (table 1). The sensitivity analysis shows limited sensitivity in the diagnostic tests in all sites led to underestimation of STD prevalence during the trials while specificity, in contrast, did not markedly influence apparent STD rates (table 4). An exception was HSV-2 infection for which the effects of sensitivity and specificity in Mwanza and Masaka almost cancelled out. The underestimation was accentuated when a screening test with limited sensitivity was used as was the case when measuring NG and CT among men in Mwanza. The selection of those reporting sexual activity for measuring HSV-2 seroprevalence in Rakai did not appreciably bias the apparent prevalence.

DISCUSSION
Previously published results from the Mwanza, Masaka, and Rakai trials suggested STD prevalences to be comparable in the three sites at baseline.[2][12][13] We have shown that, because of diagnostic errors and selection bias, the levels of curable STD may have been underestimated. This was especially the case in Mwanza, and taking these factors into account, prevalences of curable STD proved to be considerably higher in Mwanza than in Masaka and Rakai for NG, CT (apart from CT in men), TV, and active, high titre syphilis. These infections of short duration reflect current sexual behaviour in the populations. On the other hand, the prevalences of HSV-2 and serological

Table 1 Published values for sensitivities and specificities of gonorrhoea, chlamydia, trichomoniasis, and HSV-2 diagnostics used in the trials

Trial and test	Sex	Gold standard	Se (%)	Sp (%)	Population	Reference
Rakai and Masaka: LCR on urine for CT	F	Gen-Probe (endocervical) /EIA (endocervical) /LCR (urine)	98.4	100	Ob-gyn	29
	F	EIA (endocervical)/ culture (endocervical) / PCR (endocervical)/ LCR (urine)	77.8	100	CSW	30
	F	EIA (endocervical)/ culture (endocervical)/ PCR (urine)/ LCR (urine)	75	100	STD clinic	31
	F	LCR (urine)/ LCR (cervical)/ DFA (cervical)	95	nr	STD clinic	32
Mean	F		86.6	100		
Rakai and Masaka: LCR on urine for CT	M	LCR (urine)/ EIA (urethral)/ PCR (urethral)/ culture (urethral)	100	100	STD clinic	30
	M	EIA (urethral)/ culture (urtheral)/ PCR (urine)/ LCR (urine)	75	100	STD clinic	31
Mean	M		87.5	100		
Mwanza: EIA on endocervical swab for CT	F	Gen-Probe (endocervical) /EIA (endocervical) /LCR (urine)	50	100	Ob-gyn	29
	F	EIA (endocervical)/ culture (endocervical) / PCR (endocervical)/ LCR (urine)	22.2	100	CSW	30
	F	EIA (endocervical)/ culture (endocervical)/ PCR (urine)/ LCR (urine)	58	100	STD clinic	31
	F	LCR (urine)/ LCR (cervical)/ DFA (cervical)	65	nr	STD clinic	32
Mean	F		48.8	100		
Mwanza: EIA on urethral swab for CT	M	LCR (urine)/ EIA (urethral)/ PCR (urethral)/ culture (urethral)	82.6	100	STD clinic	30
	M	EIA (urethral)/ culture (urethral)/ PCR (urine)/ LCR (urine)	58	100	STD clinic	31
Mean	M		70.3	100		
Rakai and Masaka: LCR on urine for NG	F	LCR (urine)/ MOMP-LCR (urine)/ culture (endocervical)	50	100	STD clinic	21
	F	LCR (urine)/ Pilin-LCR (urine)/ Culture (endocervical)	94.7	100	STD clinic	33
	F	Culture (endocervical)/ Pilin and protein-1 LCR (urine)	91.7	100	STD clinic	34
	F	Pilin and protein1-LCR (urine)/ LCR (endocervical)/ culture (endocervical)	94.6	100	STD clinic	35
	F	PCR (urine)/ LCR (urine)/ culture (endocervical)	88.2	100	Adolescent clinic	36
Mean	F		83.8	100		
Rakai and Masaka: LCR on urine for NG	M	LCR (urine)/ MOMP-LCR (urine)/ culture (urethral)	88.9	100	STD clinic	21
Mwanza: culture on endocervical swab for NG	F	LCR (urine)/ MOMP-LCR (urine)/ Culture (endocervical)	50	100	STD clinic	21
	F	LCR (urine)/ Pilin-LCR (urine)/ Culture (endocervical)	84.2	100	STD clinic	33
	F	Culture (endocervical)/ Pilin and protein-1 LCR (urine)	58.3	100	STD clinic	34
	F	Pilin and protein1-LCR (urine)/ LCR (endocervical)/ culture (endocervical)	96.4	100	STD clinic	35
	F	PCR (urine)/ LCR (urine)/ culture (endocervical)	82.3	98.9	Adolescent clinic	36
Mean	F		74.2	99.8		
Mwanza: Gram stain on urethral smear for NG	M	LCR (urine)/ MOMP-LCR (urine)/ culture (urethral)	83.3	nr*	STD clinic	21
Rakai: InPouchTV on vaginal swab for TV	F	InPouch Culture (vaginal)/ Diamond's culture (vaginal)/ wet mount (vaginal)	88.3	100	ANC	37
	F	PCR (vaginal)/ InPouchTV culture (vaginal)/ wet mount (vaginal)	69.7	100	Army medical clinic	38
Mean	F		79	100		
Mwanza: Wet mount microscopy on vaginal swab for TV	F	InPouch Culture (vaginal)/ Diamond's culture (vaginal)/ wet mount (vaginal)	85.3	99.5	ANC	37
	F	PCR (vaginal)/ InPouchTV culture (vaginal)/ wet mount (vaginal)	36.4	100	Army medical clinic	38
Mean	F		60.9	99.8		
Rakai: CDC assay for HSV-2	F/M	Western blot	92	100		16
Mwanza and Masaka: Monoclonal antibody blocking assay for HSV-2	F/M	Western blot	91	93	Rural African population	17

*The specificity was not reported. We assumed it was 100%.
F = female, M = male, NG = gonorrhoea, CT = chlamydia, TV = trichomoniasis, HSV-2 = herpes simplex virus type 2, LCR = ligase chain reaction, PCR = polymerase chain reaction, EIA = enzyme immunoassay; MOMP = major outer membrane protein; ANC = antenatal clinic, Se = sensitivity, Sp = specificity, Ob-gyn = obstetrical/gynaecology patients, CSW = commercial sex workers, nr = not reported.

Table 2 Diagnostic test, screening test and selection bias adjustments for STD measured at basline in Rakai (Rk), Masaka (Mk), and Mwanza (Mw)

	Diagnostic test adjustment		Screening test adjustment		
	Se (range)	Sp (range)	Se′	p_{screen}	Selection bias adjustment
Rk, CT	F: 86.6 (75–98.4) M: 87.5 (75–100)	F: 100 (na) M: 100 (na)	na	na	na
Rk, NG	F: 83.8 (50–94.7) M: 88.9 (na)	F: 100 (na) M: 100 (na)	na	na	na
Rk, TV	F: 79.0 (69.7–88.3)	F: 100 (na)	na	na	na
Rk, HSV-2	92 (na)	100 (na)	na	na	% sexually active 15–19 years F: 0.71 20–24 years F: 0.99 25–29 years F: 1 15–19 years M: 0.52 20–24 years M: 0.94 25–29 years M: 1
Mk, CT	F: 86.6 (75–98.4) M: 87.5 (75–100)	F: 100 (na) M: 100 (na)	na	na	na
Mk, NG	F: 83.8 (50–94.7) M: 88.9 (na)	F: 100 (na) M: 100 (na)	na	na	na
Mk, HSV-2	91 (na)	93 (na)	na		na
Mw, CT	F: 48.8 (22.2–65) M: 70.3 (58–82.6)	F: 100 (na) M: 100 (na)	F: na M: 53.3	F: na M: 0.25	na
Mw, NG	F: 74.2 (50–96.4) M: 83.3 (na)	F: 99.8 (98.9–100) M: 100 (na)	F: na M: 87.5	F: na M: 0.25	na
Mw, TV	F: 60.9 (36.4–85.3)	F: 99.8 (99.5–100)	na		na
Mw, HSV-2	91 (na)	93 (na)	na		na

Observed prevalence adjusted according to equation 1 if no screening test used, according to equation 2 when screening test used. For selection bias adjustments see text for details. M = male, F = female, NG = gonorrhoea, CT = Chlamydia, TV = trichomoniasis, HSV-2 = herpes simplex virus type 2, Se = sensitivity, Sp = specificity, Se′ = sensitivity of screening test, p_{screen} = prevalence of positive screening test, na = not applicable.

Table 3 Observed and adjusted prevalence among women and men for CT, NG, TV, HSV-2 serology, serological syphilis (TPHA+/RPR+) and active high titre syphilis (TPHA+/RPR≥1:8) in Rakai, Masaka, and Mwanza

	Observed prevalence (%) and 95% CI			Adjusted prevalence (%) and range		
	Rakai	Masaka	Mwanza	Rakai	Masaka	Mwanza
Women						
CT (15–39 years)	2.7 (2.1 to 3.4)	1.4 (0.7 to 2.1)	6.4 (4.8 to 7.9)*	3.2 (2.8–3.7)	1.6 (1.4–1.8)	13.0 (9.8–28.6)
NG (15–39 years)	1.6 (1.1 to 2.1)	1.5 (0.7 to 2.2)	1.9 (1.0 to 2.8)*	1.9 (1.7–3.3)	1.8 (1.6–3.0)	2.3 (0.8–3.8)
TV (15–49 years)	24.3 (23.2 to 25.4)	–	25.6 (23.1 to 28.1)*	30.8 (27.6–34.9)	–	41.9 (29.7–70.5)
HSV-2 (15–29 years)	41.9 (36.8 to 47.0)	44.1 (42.4 to 45.8)	46.8 (40.2 to 53.4)	42.8	44.2	47.4
Serological syphilis (15–54 years)	9.9 (9.2 to 10.7)	11.3 (10.5 to 12.1)	8.9 (8.2 to 9.6)	nar	nar	nar
High titre syphilis (15–54 years)	1.4 (1.1 to 1.7)	0.7 (0.5 to 0.9)	6.3 (5.8 to 6.9)	nar	nar	nar
Men						
CT (15–39 years)	2.4 (1.7 to 3.0)	1.9 (0.9 to 2.9)	1.0 (0.7 to 1.2)	2.7 (2.4–3.1)	2.2 (1.9–2.5)	2.3 (1.9–2.8)
NG (15–39 years)	0.9 (0.5 to 1.4)	0.8 (0.1 to 1.4)	2.1 (1.6 to 2.5)	1.1	0.9	2.8
HSV-2 (15–29 years)	23.0 (18.4 to 27.6)	20.2 (18.6 to 21.6)	18.2 (12.8 to 23.6)	21	16.8	13.3
Serological syphilis (15–54 years)	9.6 (8.8 to 10.4)	11.3 (10.5 to 12.3)	7.3 (6.6 to 8.0)	nar	nar	nar
High-titre syphilis (15–54 years)	2.3 (1.9 to 2.7)	1.2 (0.9 to 1.5)	5.6 (5.0 to 6.2)	nar	nar	nar

*Prevalences are for antenatal clinic women.
Prevalences are age standardised by 5 year age groups according to average age structure of Rakai district, Masaka district, and rural Mwanza Region.[7][8] NG = gonorrhoea; CT = chlamydia; TV = trichomoniasis; HSV-2 = herpes simplex virus type 2, nar= no adjustment required.

syphilis tended to be similar or higher in the Ugandan sites compared to Mwanza but these STD markers reflect long term sexual behaviour in the population. The higher prevalence of short duration STD in Mwanza compared to the Ugandan sites may be the result of higher risk sexual behaviour in Mwanza relative to that in Masaka and Rakai at the baseline of the trials.[22] The similar or higher prevalence of long duration STD markers (HSV-2 and serological syphilis) in the Ugandan sites compared to Mwanza may be explained by risk behaviour having been higher there in the past compared to that measured in Mwanza.[23–26] This is supported by the higher prevalence of HIV in Rakai (16%) and Masaka (12%) compared to Mwanza (4%). The outbreak of civil unrest in Uganda and subsequent stabilisation may help explain the observed patterns of HIV prevalence and sexual behaviour in the Ugandan sites at the baseline of the trials. The natural dynamics of HIV epidemics may also have a role.

Several limitations in our analysis must be highlighted. Regarding selection bias, we must consider selection of the ANC sample in Mwanza compared to women in Rakai and Masaka when comparing NG, CT, and TV prevalences. Our results indicated higher STD prevalence among ANC women in Mwanza compared to all women in Rakai and Masaka. About 90% of pregnant women in Mwanza attend ANC so we would not expect selective factors such as STD symptoms to bias STD prevalence in the ANC sample upward.[27] As previously discussed it seems most likely the ANC sample in Mwanza would have underestimated STD prevalence in the general population.[19][20] Had we made adjustments for this the STD prevalence in women in Mwanza would have been even higher, further accentuating the difference between sites. Hence, our finding of higher STD prevalence among women in Mwanza was most probably not affected by our decision to ignore selection bias in this sample.

When comparing NG and CT prevalence in men we have to consider limitations of the performance of the complicated screening algorithm used in Mwanza. We inferred the performance of the screening algorithm (urine LED tests and

Table 4 Sensitivity analysis for STD prevalence adjustments for gonorrhoea (NG) in males and HSV-2 in females

	NG males			HSV-2 females		
	Rakai	Masaka	Mwanza	Rakai	Masaka	Mwanza
Observed prevalence (%)	0.93	0.80	2.05	41.9	44.1	46.8
Prevalence after adjustment for:						
Se alone	1.05	0.90	2.46	45.5	48.5	51.4
Sp alone	0.93	0.80	2.05	41.9	39.9	42.8
Se + Sp	**1.05**	**0.90**	2.46	45.5	**44.2**	**47.4**
Se' alone	na	na	2.34	na	na	na
Se' UL alone	na	na	2.16	na	na	na
Se' LL alone	na	na	4.66	na	na	na
Se+Sp+Se'	na	na	**2.81**	na	na	na
Se+Sp+Se'UL	na	na	2.59	na	na	na
Se+Sp+Se'LL	na	na	5.59	na	na	na
Selection bias UL	na	na	na	39.4	na	na
Selection bias	na	na	na	36.9	na	na
Se+Sp+selection bias UL	na	na	na	44.2	na	na
Se+Sp+selection bias	na	na	na	**42.8**	na	na

Highlighted values correspond to values from table 3. Se = sensitivity, Sp = specificity, Se' = sensitivity of screening test, p_{screen} = prevalence of positive screening test, UL = upper limit, LL = lower limit, na= not applicable.

reporting of symptoms or signs on clinical examination) from its use in a general population survey among men in a rural community of Mwanza region, Misungwi, in 1996.[11] In Misungwi, the performance of the screening algorithm was compared to LCR on urine for CT and culture on urethral swabs for NG. This inference was appropriate because the proportion of men positive for the LED test, reporting symptoms, or with signs on clinical examination were known to be similar between the Misungwi study and the Mwanza trial. In the trial, the distribution of reasons for positivity on the screening test was 92% (out of 1451 men) LED positive only, 4% complaining of discharge or ulcerative symptoms or had signs on clinical examination but LED negative, and 4% with symptoms or signs and LED positive.[28] In the Misungwi study of 438 men, 96% who tested positive on the screening algorithm had a positive LED test only, 2% had signs or symptoms only, and 2% had symptoms or signs and were LED positive. Furthermore, if the sensitivity of the screening algorithm differed between the two populations, our adjusted prevalence estimates for the trial population would be biased but this is probably not the case since the determinants of sensitivity, such as the prevalence of schistosomiasis, are similar in Misungwi and the Mwanza trial communities.

Finally, limitations in the performance of the diagnostic tests used influenced the measured prevalence in the two trials. Owing to the earlier date of the Mwanza trial, it was not feasible to use highly sensitive tests which were not available or affordable at that time. This explains why diagnostic biases were larger in Mwanza than in Masaka and Rakai. In this analysis, low sensitivity had a more critical role than low specificity and most adjustments of prevalences were upward (table 3) since the specificities were high (table 1). For NG and CT in men in Mwanza, the effect of low sensitivity of the diagnostic test was enhanced by the low sensitivity of the screening test. The true magnitude of these diagnostic biases are however not certain, since we used estimates of test sensitivity and specificity mostly from laboratory evaluations in Western settings (table 1), where sensitivity and/or specificity may be better than under conditions in the rural African sites. It is unclear how much diagnostic test performance may have declined in the trial settings. In addition, the majority of estimates for sensitivity and specificity of diagnostic tests came from symptomatic (STD clinic) populations. We might, however, expect the sensitivity of the tests to be worse for asymptomatic individuals compared to symptomatic individuals and this would lead to

underestimation of STD prevalence in all sites. Finally, the gold standards used to measure sensitivity and specificity may be imperfect which would limit our ability to estimate true absolute prevalence. However, the same gold standards were used for the different tests so our site comparison of the relative prevalences in the three populations should be valid even if the absolute prevalences are different owing to an imperfect gold standard.

This illustration of the importance of sample selection and diagnostic techniques in biasing observed STD prevalences has implications for the conduct and interpretation of randomised trials and other surveys with STD prevalence outcomes. The complexity of the adjustments made for the screening algorithm in Mwanza indicates it is preferable wherever possible to use random subsamples of the general population to monitor STD prevalence.

With respect to diagnostic biases in intervention trials, not only their effects on baseline prevalence but also the observed impact of the intervention on STD prevalences must be considered. Limitations in accuracy of diagnostic tests will in general result in dilution of observed impact on prevalence (that is, differences between arms at follow up or reductions over time in one arm). Limited specificity is generally of more concern in this respect than limited sensitivity. As an example, assume the *observed* impact of an intervention on CT was to reduce prevalence from 4% to 2% (relative risk = 0.5, absolute difference = −2%). With 100% sensitivity and 99% specificity, the true prevalence—after adjustment for imperfect specificity—would be 3% in the comparison arm and 1% in the intervention arm, consistent with a true relative risk of 0.33. If sensitivity is limited (90%) but specificity is 100%, the relative difference in the true prevalence between arms or time points is the same as the relative difference in observed prevalence (0.5) but the absolute difference is larger, −2.2%. If both are limited by these assumed amounts, the true relative risk is 0.33 and the true absolute difference is −2.25%.

In conclusion, the biases in STD prevalences in the general population cohorts of the three trials, and particularly in Mwanza, illustrate the critical attention which should be paid to selecting the population in which the STD is to be measured and the diagnostic technique to be used when conducting a research study with STD prevalence outcomes. Notably, besides the properties of diagnostic tests, it is preferable to screen whole populations or random sample, and to avoid screening algorithms based on symptomatology which omit asymptomatic infections. In intervention trials, high specificity is critical to avoid dilution of impact measures. Our

adjustment methods suggest that, after taking into account diagnostic error and selection bias, the prevalences of curable STD were higher in Mwanza than in Masaka and Rakai at the baseline of the trials. Based on this, we might expect STD treatment to have a larger impact on HIV incidence in Mwanza than in the Ugandan sites. This finding may, in part, explain why STD treatment was associated with lower HIV incidence in Mwanza but not in the Ugandan sites.

ACKNOWLEDGEMENTS

CONTRIBUTORS

This study was designed, implemented, and analysed by KO, EK, and RW. It is part of a larger project which was planned, designed, implemented, and supervised by HG and RH and involved close collaboration with RG, AK, DS, SV, and MW. The collection, analysis, and use of trial data was conducted by PH, AK and AO for the Masaka data, RG, DS, and MW for the Rakai data and HG, RH, and JC for the Mwanza data. The manuscript was prepared by KO, EK, HG, and RH with contributions from all other authors.

APPENDIX 1
Derivation of adjustment for diagnostics and screening, equation 2

Test 1 Screening test

	True status		
Test	Positive	Negative	Total
Positive	a	b	e
Negative	c	d	f
Total	g	h	N

Test 2 Diagnostic test

	True status		
Test	Positive	Negative	Total
Positive	k	l	i
Negative	m	n	j
Total	a	b	e

Let:

Se = sensitivity of diagnostic test, Sp = specificity of diagnostic test, Se′ = sensitivity of screening test, Sp′ = specificity of screening test.

Then expected values of cell frequencies are as follows:

$a = Np_{true}Se'$ $b = N(1 - p_{true})(1 - Sp')$
Hence $p_{screen} = (a + b)/N = p_{true}Se' + (1 - p_{true})(1 - Sp')$
$k = Np_{true}Se'Se$ $l = N(1 - p_{true})(1 - Sp')(1 - Sp)$
Hence $p_{observed} = (k + l)/N = p_{true}Se'Se + (1 - p_{true})(1 - Sp')(1 - Sp)$
$= p_{true}Se'Se + (p_{screen} - p_{true}Se')(1 - Sp)$
$= p_{true}(Se'Se - Se' + SpSe') + p_{screen}(1 - Sp)$

This yields equation 2 in Methods:

$$p_{true} = \frac{[p_{observed} - p_{screen}(1 - Sp)]}{[Se'(Se + Sp - 1)]}$$

....

Authors' affiliations
K K Orroth, R G White, R J Hayes, H Grosskurth, London School of Hygiene and Tropical Medicine, London, UK
E L Korenromp, S J de Vlas, Department of Public Health, Erasmus University, Rotterdam, Netherlands
J Changalucha, National Institute for Medical Research, Mwanza, Tanzania
R H Gray, Johns Hopkins School of Public Health, Baltimore, USA
P Hughes, A Kamali, A Ojwiya, MRC Programme on AIDS in Uganda, Entebbe, Uganda
D Serwadda, Institute of Public Health, Faculty of Medicine, Makerere University, Kampala, Uganda
M J Wawer, Columbia University School of Public Health, New York, USA

REFERENCES

1 **Grosskurth H**, Mosha F, Todd J, et al. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomized controlled trial. Lancet 1995;**346**:530–6.
2 **Kamali A**, Kintu P, Nalweyiso N, et al. A randomised controlled trial (RCT) of syndromic STD management and behavioural change interventions to reduce HIV infection in rural Uganda. Int J STD AIDS 2001;**12(Suppl 2)**:46.
3 **Wawer MJ**, Sewankambo NK, Serwadda D, et al. Control of sexually transmitted diseases for AIDS prevention in Uganda: a randomised community trial. Lancet 1999;**353**:525–35.
4 **Hayes R**, Grosskurth H, ka-Gina G. Impact of improved treatment of sexually transmitted disease on HIV infection: authors' reply. Lancet 1995;**346**:1159–60.
5 **Kamali A**, Quigley M, Nakiyingi J, et al. A community randomised trial of sexual behaviour and syndromic STI management interventions on HIV-1 transmission in rural Uganda. (Submitted.)
6 **Grosskurth H**, Gray RH, Hayes R, et al. Control of sexually transmitted diseases for HIV-1 prevention: understanding the implications of the Mwanza and Rakai trials. Lancet 2000;**355**:1981–7.
7 **Statistics Bo**. 1988 Population Census: National Profile, Fertility and Mortality Data. Dar es Salaam: President's Office, Planning Commission; 1994 September, 1994.
8 **Republic of Uganda**. 1991 Population and Housing Census, Rakai district. Entebbe: Statistical division, Ministry of Planning 1991.
9 **Kelsey J**, Thompson WD, Evans AS. Methods in observational epidemiology. Oxford: Oxford University Press, 1986.
10 **Robinson A**, Ridgway G. Modern diagnosis and management of genital Chlamydia trachomatis infection. Br J Hosp Med 1996;**55**:388–93.
11 **Watson-Jones D**, Mugeye K, Mayaud P, et al. High prevalence of trichomoniasis in rural men in Mwanza, Tanzania: results from a population-based study. Sex Transm Infect 2000;**76**:355–62.
12 **Wawer MJ**, Gray RH, Sewankambo NK, et al. A randomized, community trial of intensive sexually transmitted disease control for AIDS prevention, Rakai, Uganda. AIDS 1998;**12**:1211–25.
13 **Grosskurth H**, Mosha F, Todd J, et al. A community trial of the impact of improved sexually transmitted disease treatment on the HIV epidemic in rural Tanzania: 2. Baseline survey results. AIDS 1995;**9**:927–34.
14 **Larsen SA**, Steiner BM, Rudolph AH. Laboratory diagnosis and interpretation of tests for syphilis. Clin Microbiol Rev 1995;**8**:1–21.
15 **Hart G**. Syphilis tests in diagnostic and therapeutic decision making. Ann Intern Med 1986;**104**:368–76.
16 **Sanchez Martinez D**, Schmid DS, Whittington W, et al. Evaluation of a test based on baculovirus-expressed glycoprotein G for detection of herpes simplex virus type-specific antibodies. J Infect Dis 1991;**164**:1196–9.
17 **Gopal R**, Gibbs T, Slomka MJ, et al. A monoclonal blocking EIA for herpes simplex virus type 2 antibody validation for seroepidemiological studies in Africa. J Virol Methods 2000;**87**:71–80.
18 **Obasi A**, Mosha F, Quigley M, et al. Antibody to HSV-2 as a marker of sexual risk behaviour in rural Mwanza. J Infect Dis 1999;**179**:16–24.
19 **Kigadye RM**, Klokke A, Nicoll A, et al. Sentinel surveillance for HIV-1 among pregnant women in a developing country: 3 years' experience and comparison with a population survey. AIDS 1993;**7**:849–55.
20 **Mosha F**, Nicoll A, Borgdorff M, et al. A population-based study of syphilis and sexually transmitted disease syndromes in north-western Tanzania. 1. Prevalence and incidence. Genitourin Med 1993;**69**:415–420.
21 **Buimer M**, Van Doornum G, Ching S, et al. Detection of chlamydia trachomatis and Neisseria gonorrhoeae by ligase chain reaction-based assays with clinical specimens from various sites: implications for diagnostic testing and screening. J Clin Microbiol 1996;**34**:2395–400.
22 **Korenromp EL**. Treatment of sexuall transmitted diseases as an HIV prevention strategy? Cofactor magnitudes, syndromic management and a reappraisal of the Mwanza and Rakai trials. Rotterdam: Erasmus University, 2001.
23 **Korenromp EL**, Bakker R, Gray RH, et al. The effect of HIV, behavioural change and STD syndromic management on STD epidemiology in sub-Saharan Africa: simulations of Uganda. Sex Transm Infect 2002;(in press.)

24 **Konde-Lule JK**, Wawer MJ, Sewankambo NK, *et al.* Adolescents, sexual behavior and HIV-1 in rural Rakai district, Uganda. *AIDS* 1997;**11**:791–9.

25 **Asiimwe-Okiror G**, Opio AA, Musinguzi J, *et al.* Change in sexual behaviour and decline in HIV infection among yount pregnant women in urban Uganda. *AIDS* 1997;**11**:1757–63.

26 **Kamali A**, Carpenter L, Whitworth J, *et al.* Seven-year trends in HIV-1 infection rates, and changes in sexual behaviour, among adults in rural Uganda. *AIDS* 2000;**14**:427–34.

27 **Tanzania Demographic and Health Survey**, 1996. Dar es Salaam, Tanzania, Calverton, Maryland: Bureau of Statistics, Planning Commission; 1997.

28 **Grosskurth H**, Mayaud P, Mosha F, *et al.* Asymptomatic gonorrhoea and chlamydial infection in rural Tanzanian men. *BMJ* 1996;**312**:277–80.

29 **Lauderdale T**, Landers L, Thorneycroft I, *et al.* Comparison of the PACE 2 assay, two amplification assays, and Clearview EIA for detection of Chlamydia trachomatis in female endocervical and urine specimens. *J Clin Microbiol* 1999;**37**:2223–9.

30 **Tan HH**, Chan RK, Teo AS, *et al.* Use of ligase chain reaction and polymerase chain reaction on urine specimens to detect Chlamydia trachomatis infections in a sexually transmitted diseases clinic in Singapore. *Ann Acad Med Singapore* 1999;**28**:245–51.

31 **Schepetiuk S**, Kok T, Martin L, *et al.* Detection of Chlamydia trachomatis in urine samples by nucleic acid tests: comparison with culture and enzyme immunoassay of genital swab specimens. *J Clin Microbiol* 1997;**35**:3355–7.

32 **Thomas BJ**, Pierpoint T, Taylor-Robinson D, *et al.* Qualitative and quantitative aspects of the ligase chain reaction assay for Chlamydia trachomatis in genital tract samples and urines. *Int J STD AIDS* 2001;**12**:589–94.

33 **Stary A**, Ching S, Teodorowicz L, *et al.* Comparison of ligase chain reaction and culture for detection of Neisseria gonorrhoeae in genital and extragenital specimens. *J Clin Microbiol* 1997;**35**:239–42.

34 **Carroll K**, Aldeen W, Morrison M, *et al.* Evaluation of the Abbott LCx ligase chain reaction assay for detection of Chlamydia trachomatis and Neisseria gonorrhoeae in urine and genital swab specimens from a sexually transmitted disease clinic population. *J Clin Microbiol* 1998;**36**:1630–3.

35 **Smith KR**, Ching S, Lee H, *et al.* Evaluation of ligase chain reaction for use with urine for identification of Neisseria gonorrhoeae in females attending a sexually transmitted disease clinic. *J Clin Microbiol* 1995;**33**:455–7.

36 **Xu K**, Glanton V, Johnson SR, *et al.* Detection of *Neisseria gonorrhoeae* infection by ligase chain reaction testing of urine among adolescent women with and without Chlamydia trachomatis infection. *Sex Transm Dis* 1998;**25**:533–8.

37 **Draper D**, Parker R, Patterson E, *et al.* Detection of Trichomonas vaginalis in pregnant women with the InPouch TV culture system. *J Clin Microbiol* 1993;**31**:1016–18.

38 **Madico G**, Quinn TC, Rompalo AM, *et al.* Diagnosis of Trichomonas vaginalis infection by PCR using vaginal swab samples. *J Clin Microbiol* 1998;**36**:3205–10.